Using k-Dominating Sets to Select Cluster Heads for Data Aggregation in Sensor Systems

Kranthi K Mamidisetty, Branden Archer, Shivakumar Sastry

The University of Akron

Overview

- Data Aggregation
- K-Domination
  - Definition
  - Embedding functions
  - Tiling in Mesh Topologies
- System Model
  - Energy Model
- Simulation Results
- Conclusions

Data Aggregation

- Resource constrained nodes
  - Limited computation and communication
  - Peer-to-peer multi-hop wireless communications
- Loosely-coupled, large-scale system
- Large volume of data, low bandwidth and less energy

Data Aggregation

- Data aggregation methods are divided in to
  - Cluster based approaches
  - Non Cluster based approaches

Data Aggregation

- Nodes collect local data and send data to the sink or monitoring station
- Amount of data handled is large and cost of communication is high because of direct communication

Data Aggregation

- Nodes in the network are grouped to communicate to a node in the network which are knows as Cluster Heads (CH)
- CH collect local data aggregate them and transmits the corresponding data to sink

Overview

- Data Aggregation
- K-Domination
  - Definition
  - Embedding functions
  - Tiling in Mesh Topologies
- System Model
  - Energy Model
- Simulation Results
- Conclusions

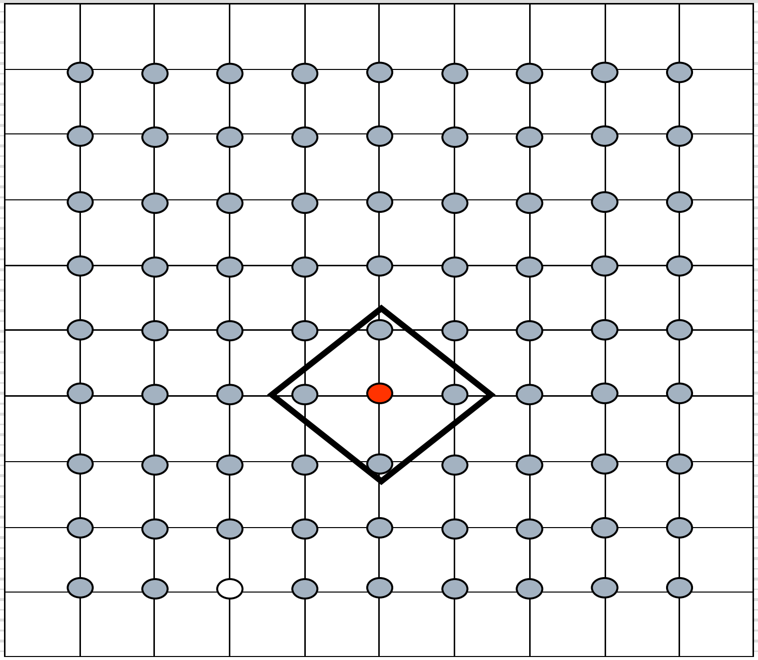
K-Domination

Definition

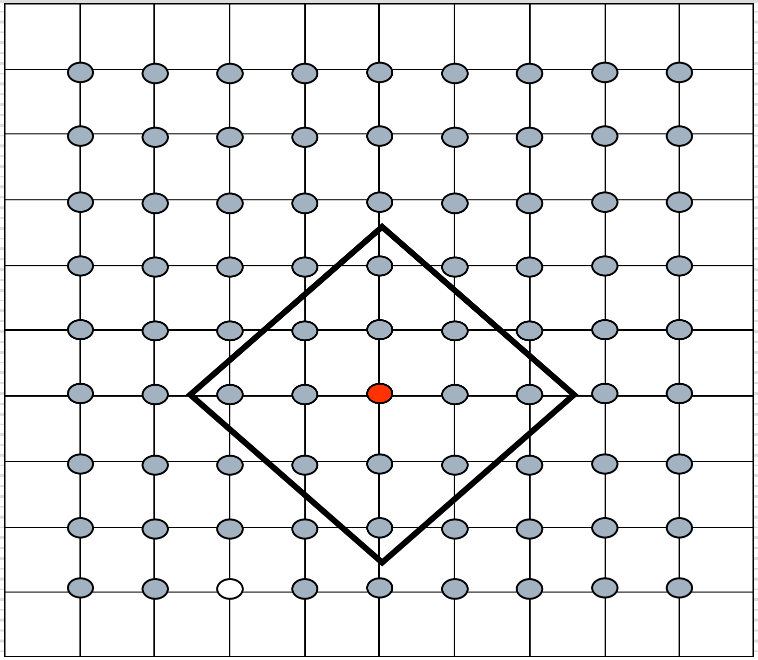
- Given a graph, G=(V,E), a subset D of V is said to be a 1-dominating set, if every node v in V is either in D or is adjacent to a node in D.
- A subset D is said to be a k-dominating set if every node v, not in D, is connected to a node in D via a path of length less than or equal to k.

K-Domination

Examples



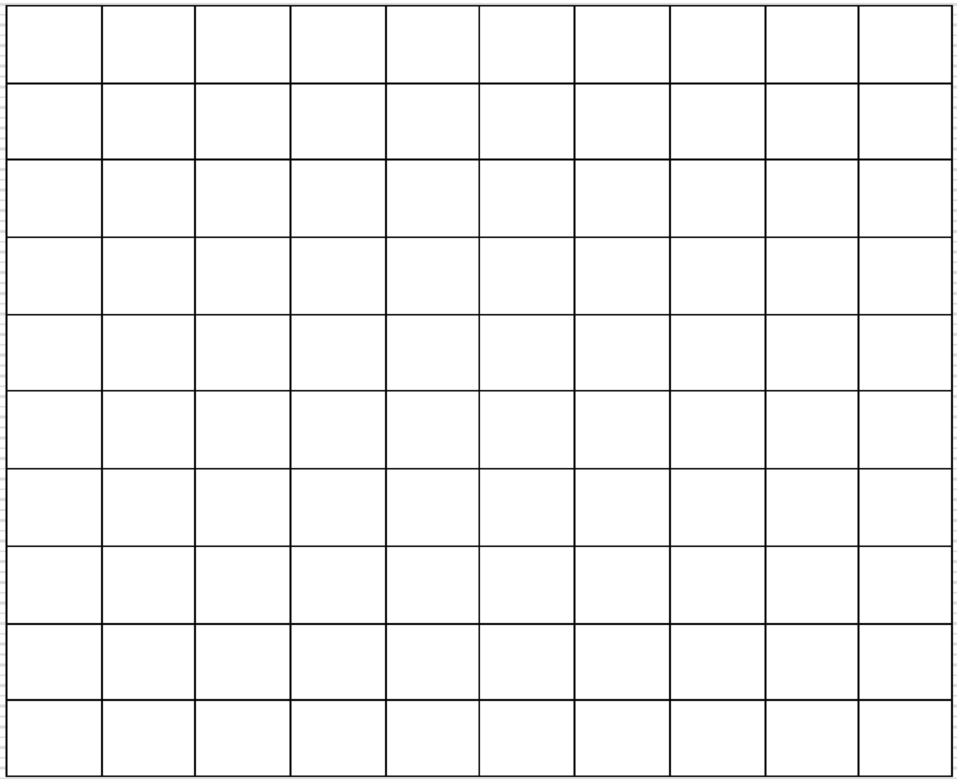
1-Domination



2-Domination

K-Domination

2D-BaseGrid



Overview

- Data Aggregation
- K-Domination
  - Definition
  - Embedding functions
  - Tiling in Mesh Topologies
- System Model
  - Energy Model
- Simulation Results
- Conclusions

K-Domination

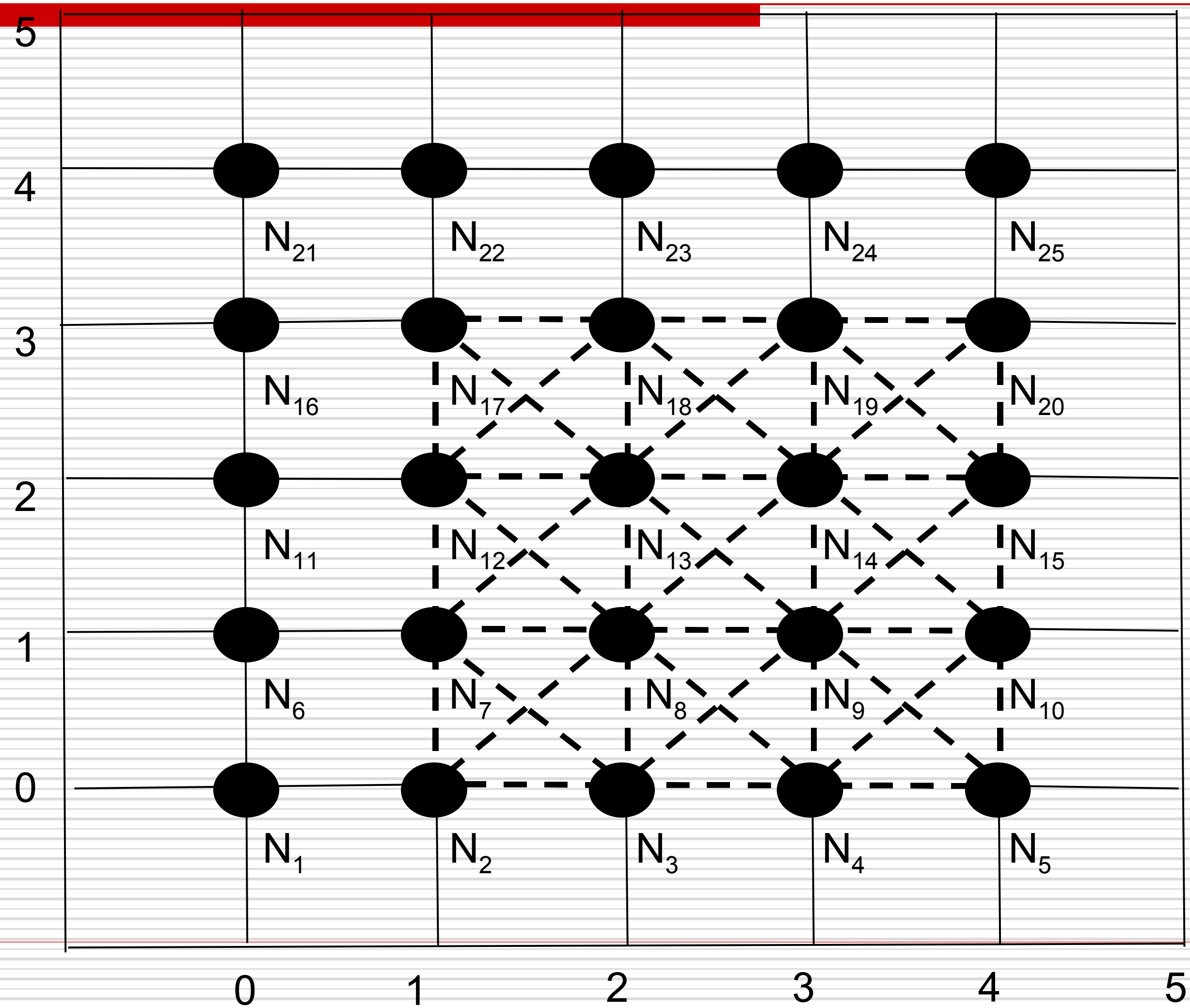
Embedding Functions

- Given a set of nodes, *Z*, an embedding function is

$$\Xi_q : Z \rightarrow N \times N \times R$$

K-Domination

Embedding Functions – 8 Neighbor Mesh Topology



K-Domination

Embedding Functions – 8 Neighbor Mesh Topology

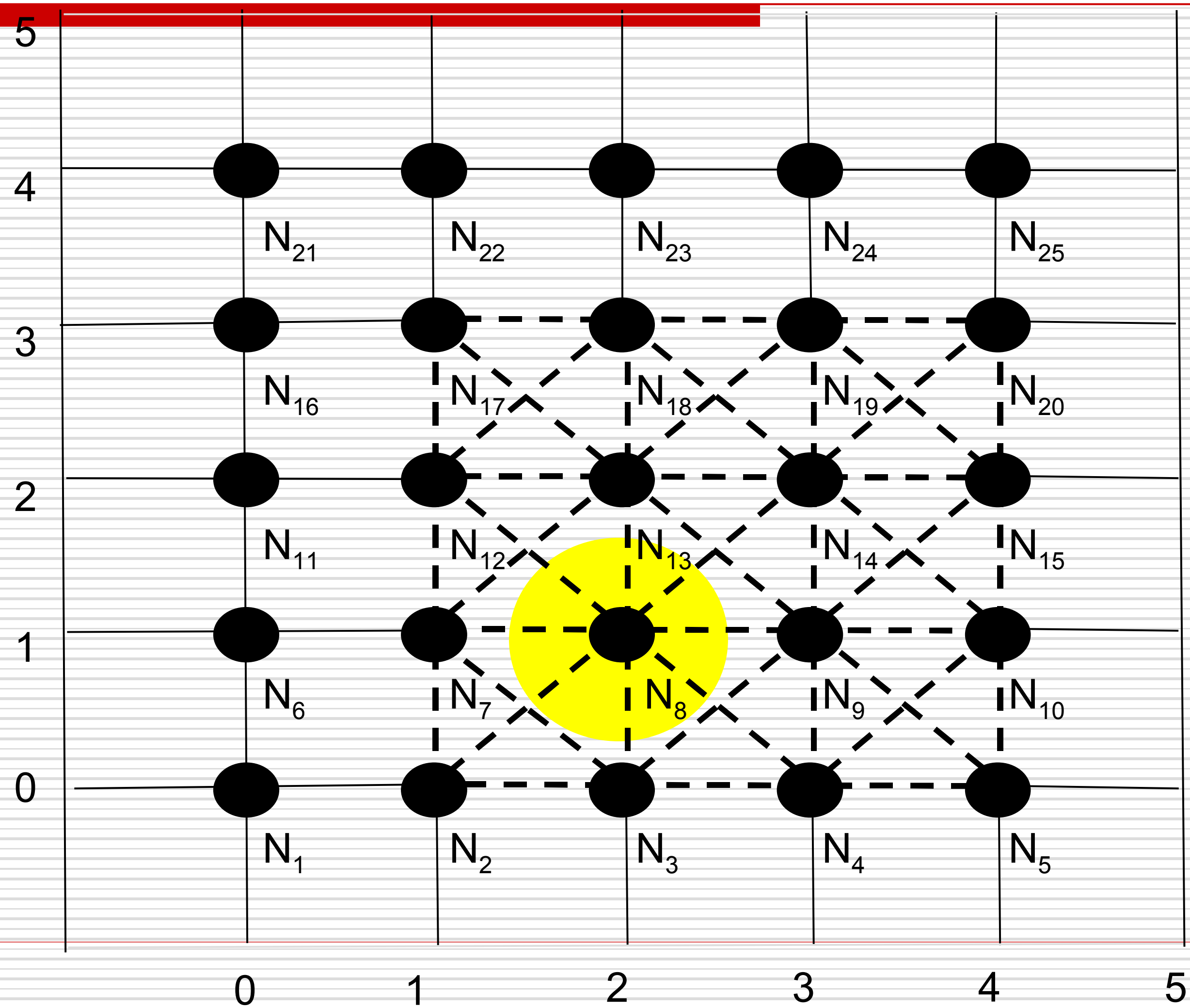
- Node Z_i is assigned to the location

$$((i-1) \bmod C, \left\lceil \frac{i}{C} \right\rceil - 1)$$

- C is number of columns.
- The transmission range is set to $\sqrt{2}$
- Example, when C = 5, node N_8 is assigned to the location (2,1).

K-Domination

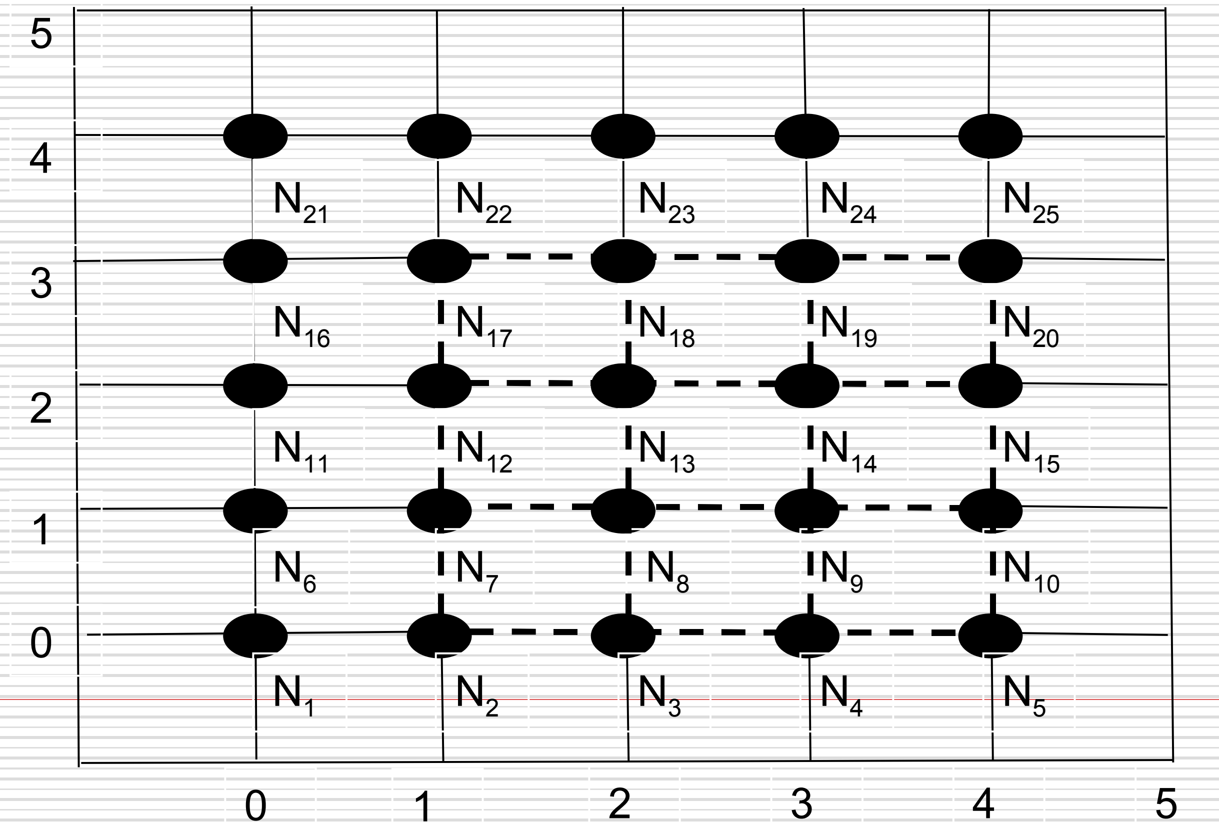
Embedding Functions – 8 Neighbor Mesh Topology



K-Domination

Embedding Functions – 4 Neighbor Mesh Topology

- Using the same location assignment with a transmission range of 1, we get the 4-Neighbor topology.



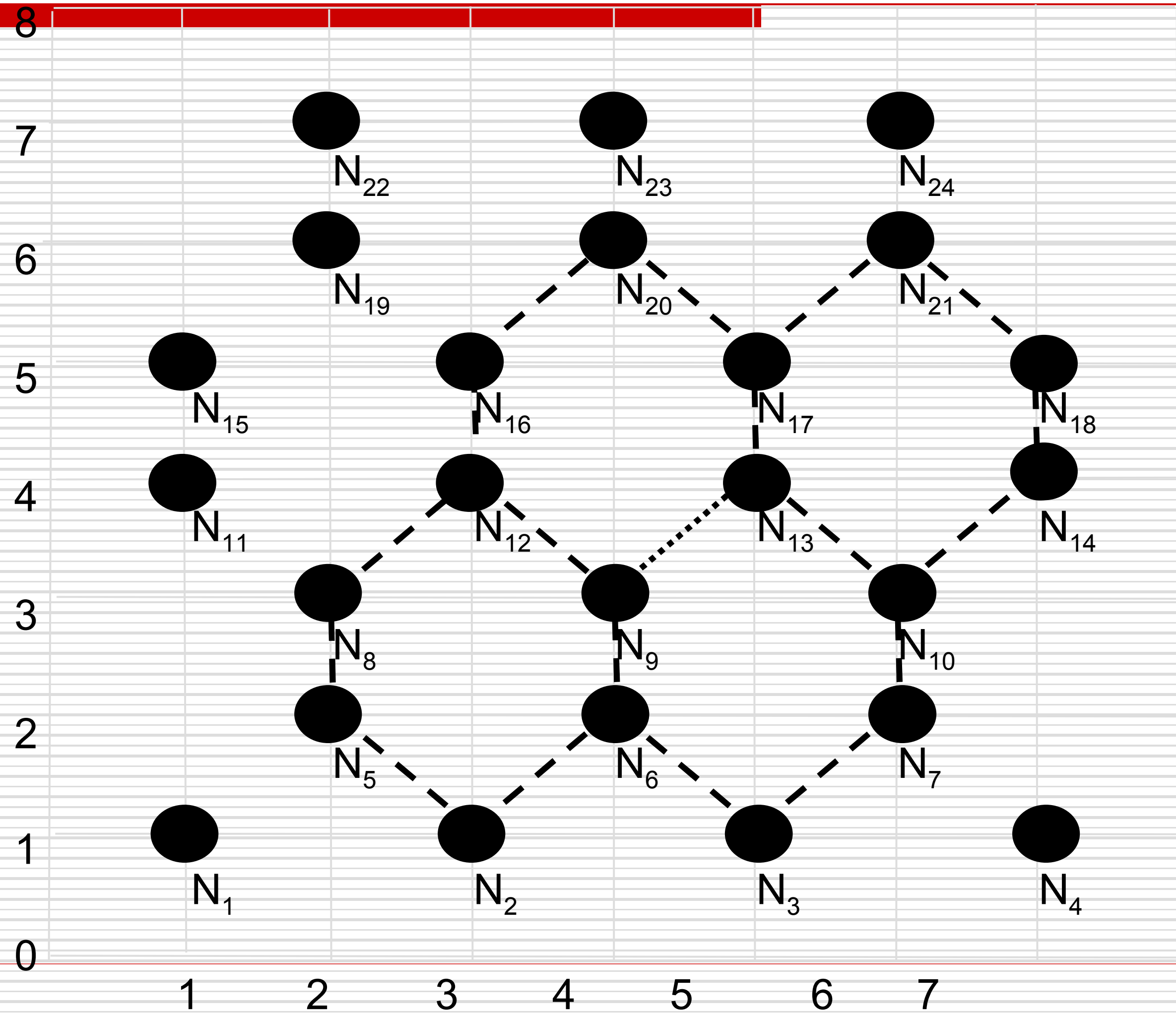
K-Domination

Other Mesh Topologies

- Two other embedding functions are for
  - 3-Neighbor and
  - 6-Neighbor

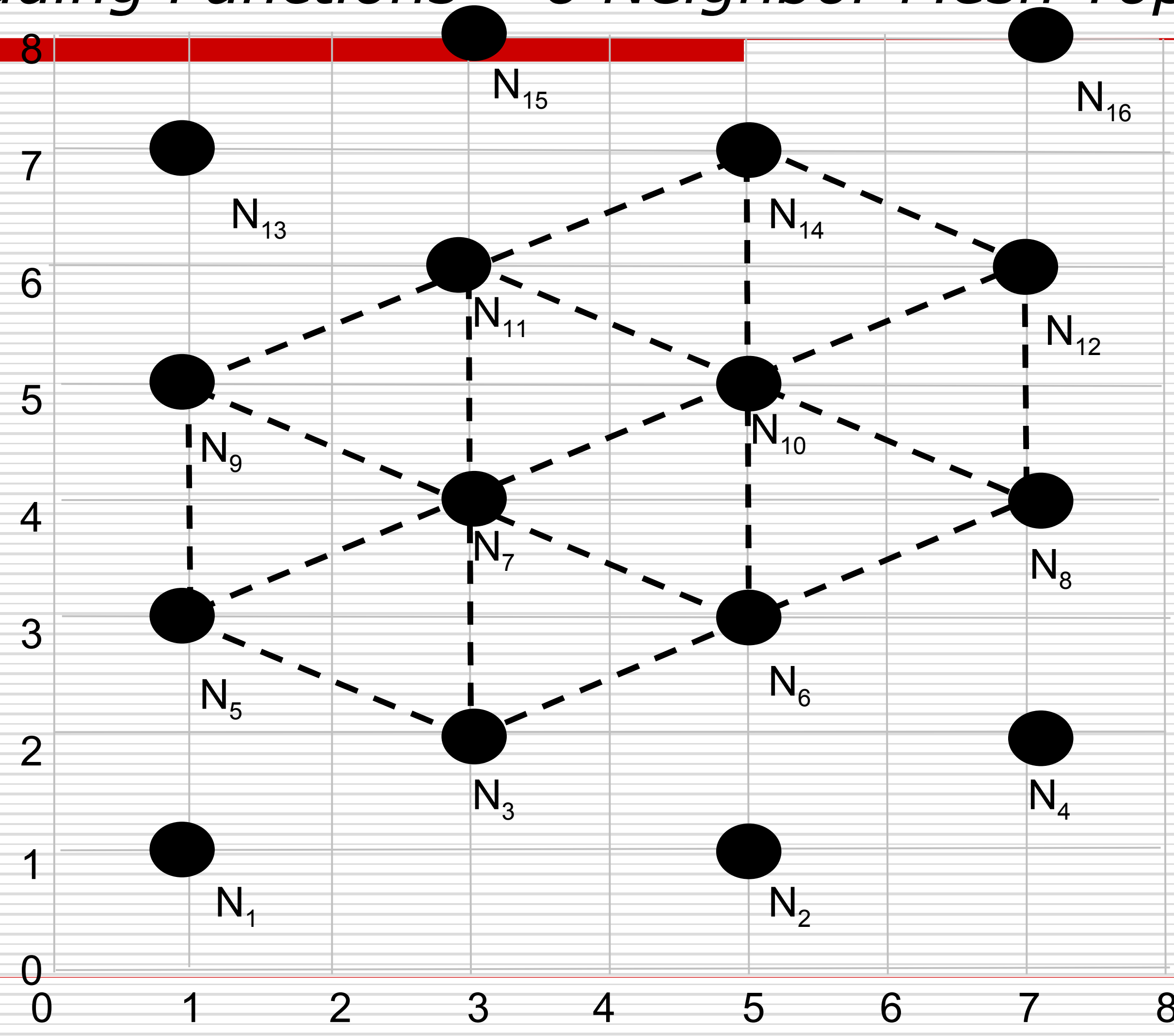
K-Domination

Embedding Functions – 3 Neighbor Mesh Topology



K-Domination

Embedding Functions – 6 Neighbor Mesh Topology



K-Domination

Embedding Functions

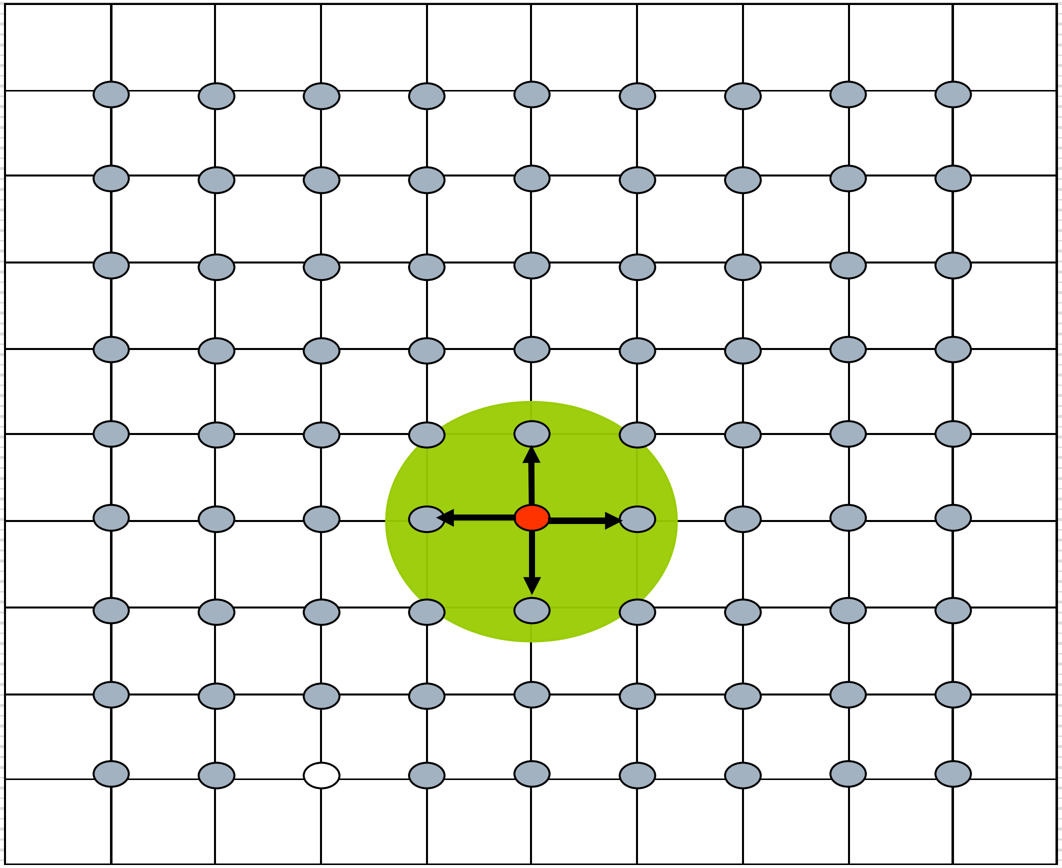
- Therefore, the notion of embedding functions helped us to view different mesh topologies in a unified manner.
- Given a set of nodes, each node is assigned to an (x,y) location on the grid with a specific transmission range.

Overview

- Data Aggregation
- K-Domination
  - Definition
  - Embedding functions
  - Tiling in Mesh Topologies
- System Model
  - Energy Model
- Simulation Results
- Conclusions

K-Domination

4 Neighbor Mesh topology



K-Domination

Tiles in Mesh topologies

- For each mesh topology there is a particular tile structure which covers the whole network.
- The center node in each tile act as CH.
- Nodes in the tile communicate to the CH.

K-Domination

Tiles in Mesh topologies

- The size of the tile and number of nodes in each tile increases as the value of 'k' in k-Domination increases
- Number of nodes in each cluster for k-Domination is $q*(k^2+k)/2$.
- q is number of neighbors for a node

K-Domination

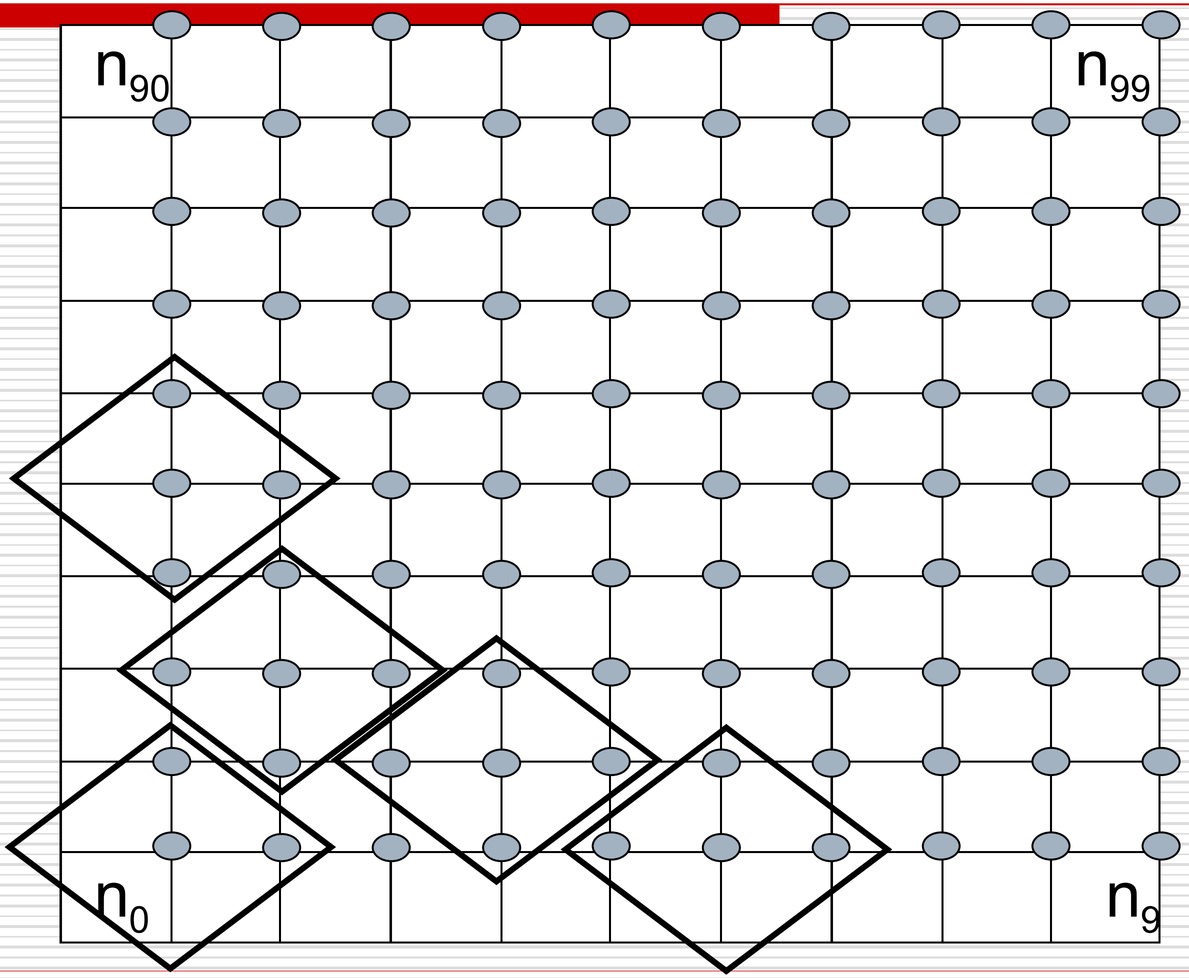
Tiles in Mesh topologies

$$\text{Number of Clusters} = \frac{\text{Number of Nodes}}{\text{Number of Nodes per cluster}}$$

- The above equation gives tentatively the number of clusters as the network size increases

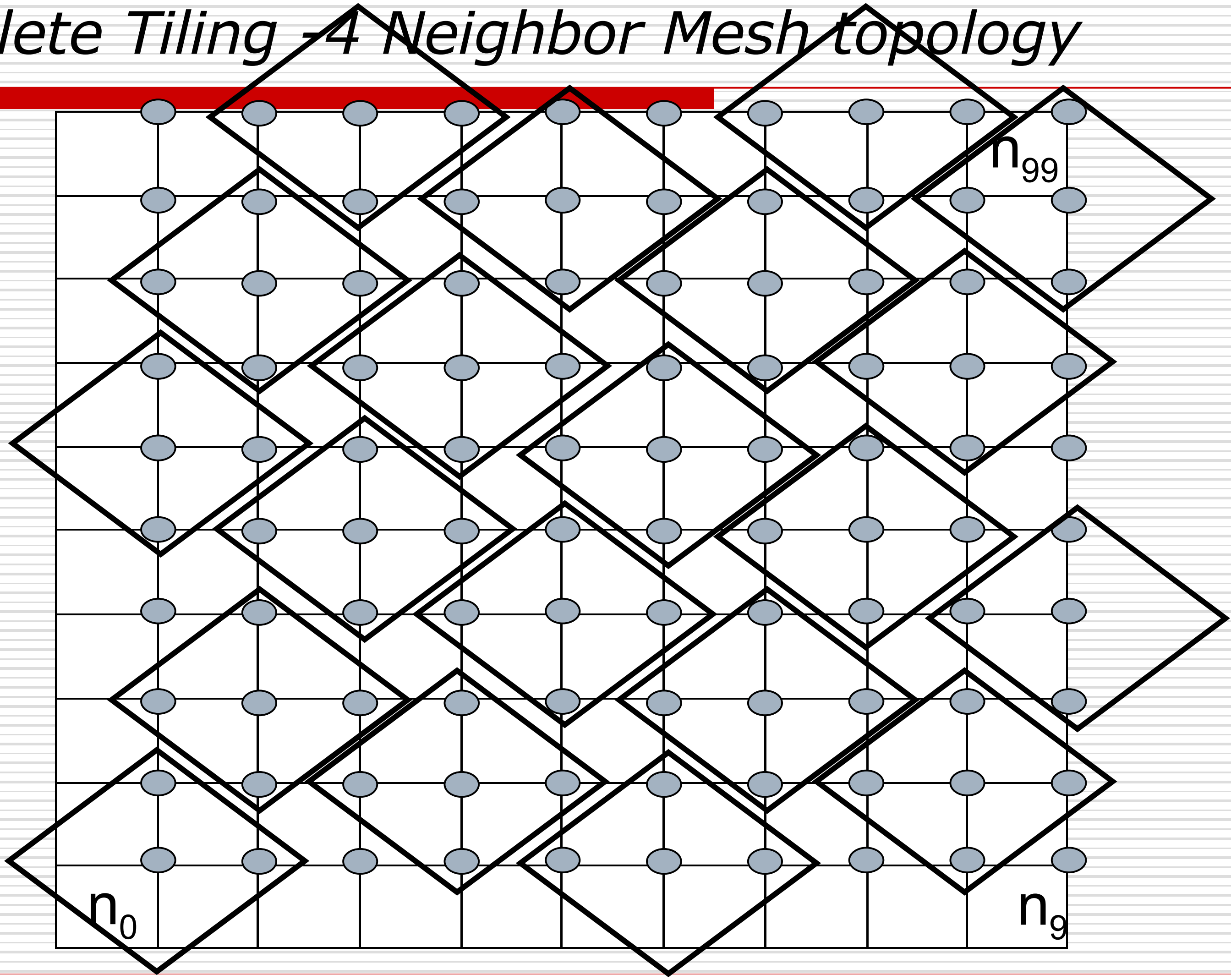
K-Domination

Example Tiling -4 Neighbor Mesh topology



K-Domination

Complete Tiling -4 Neighbor Mesh topology



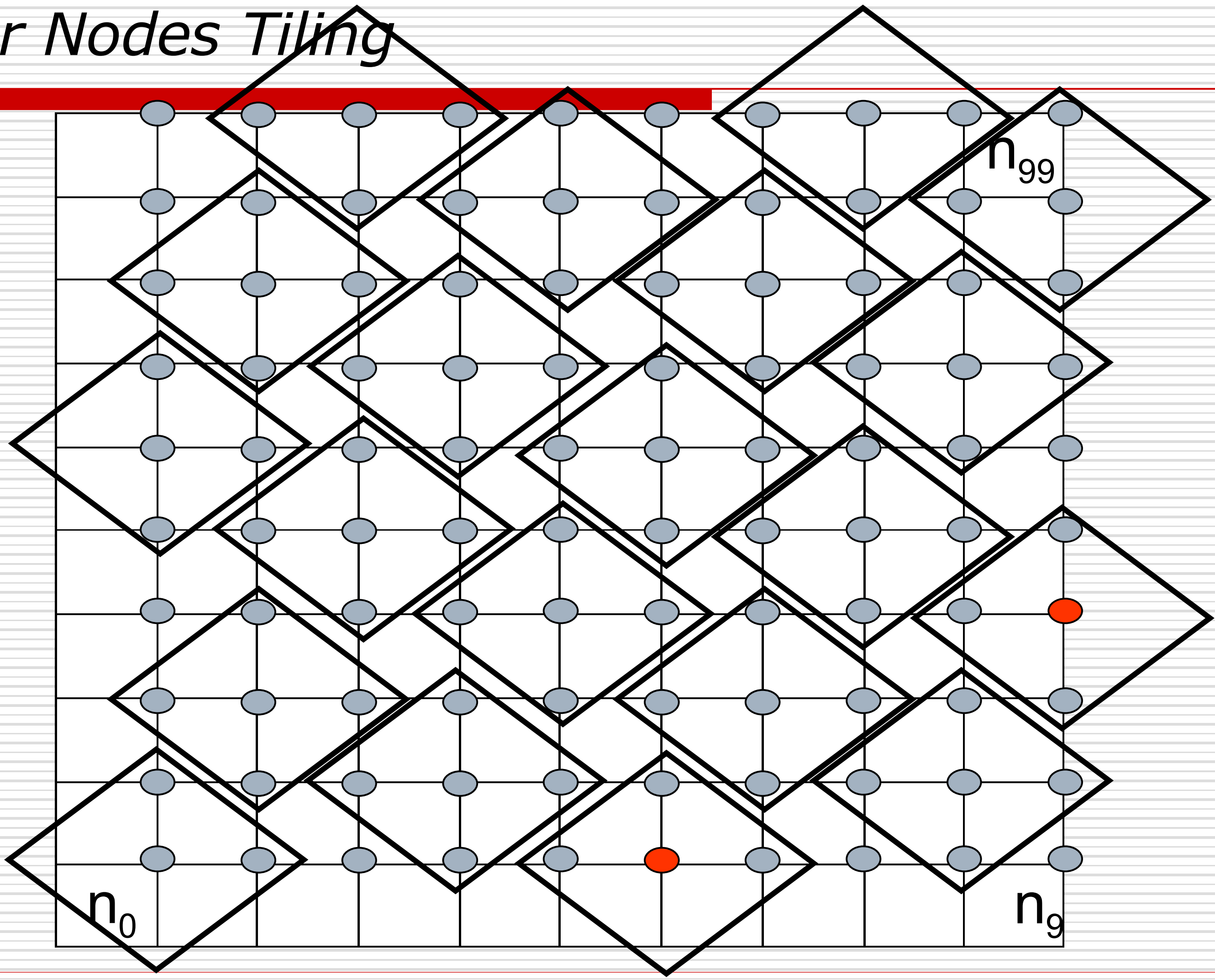
K-Domination

Border Nodes Tiling

- Some nodes along the border cannot be covered by complete tiles
- These nodes are covered by ‘subtiles’

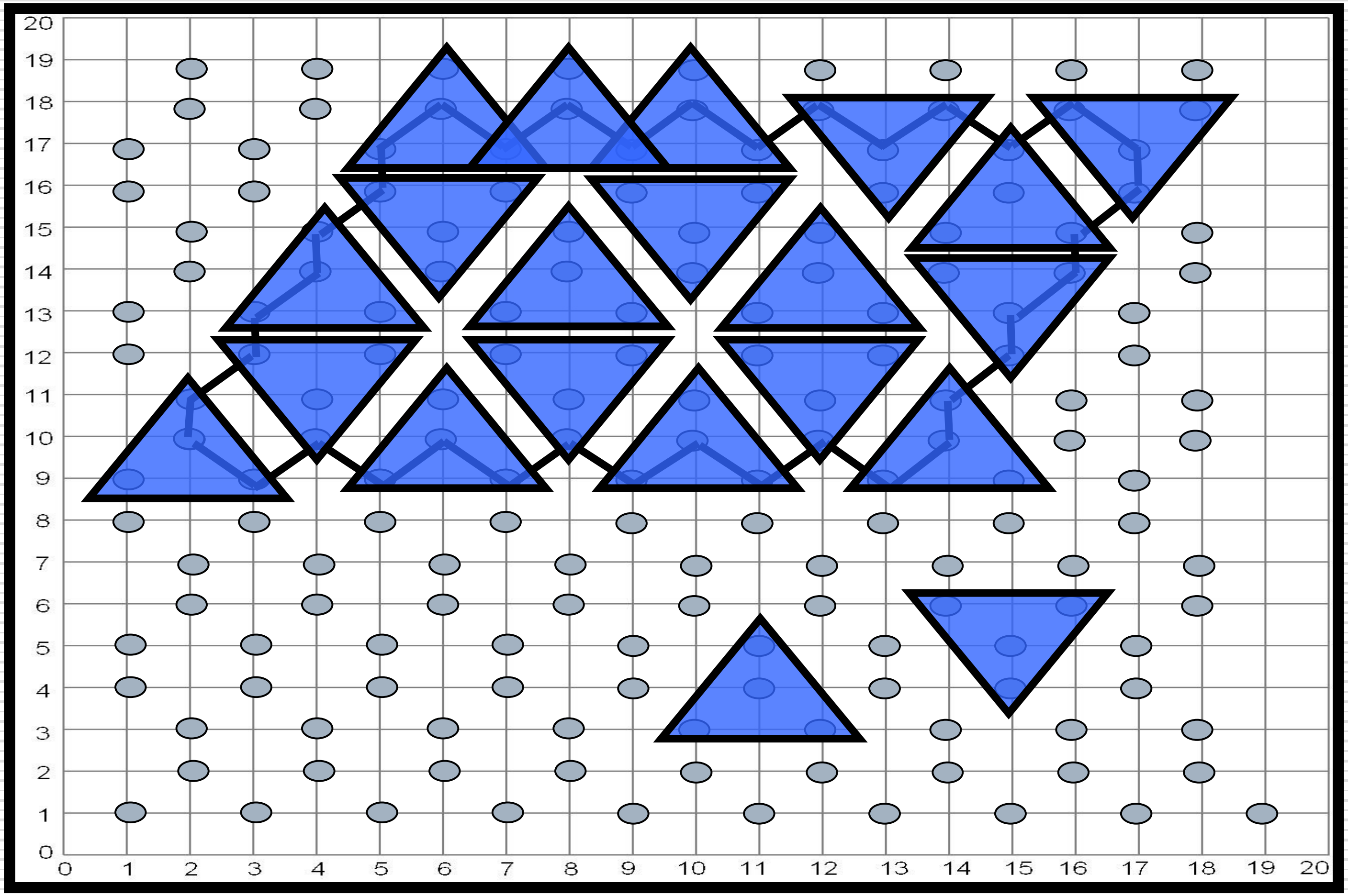
K-Domination

Border Nodes Tiling



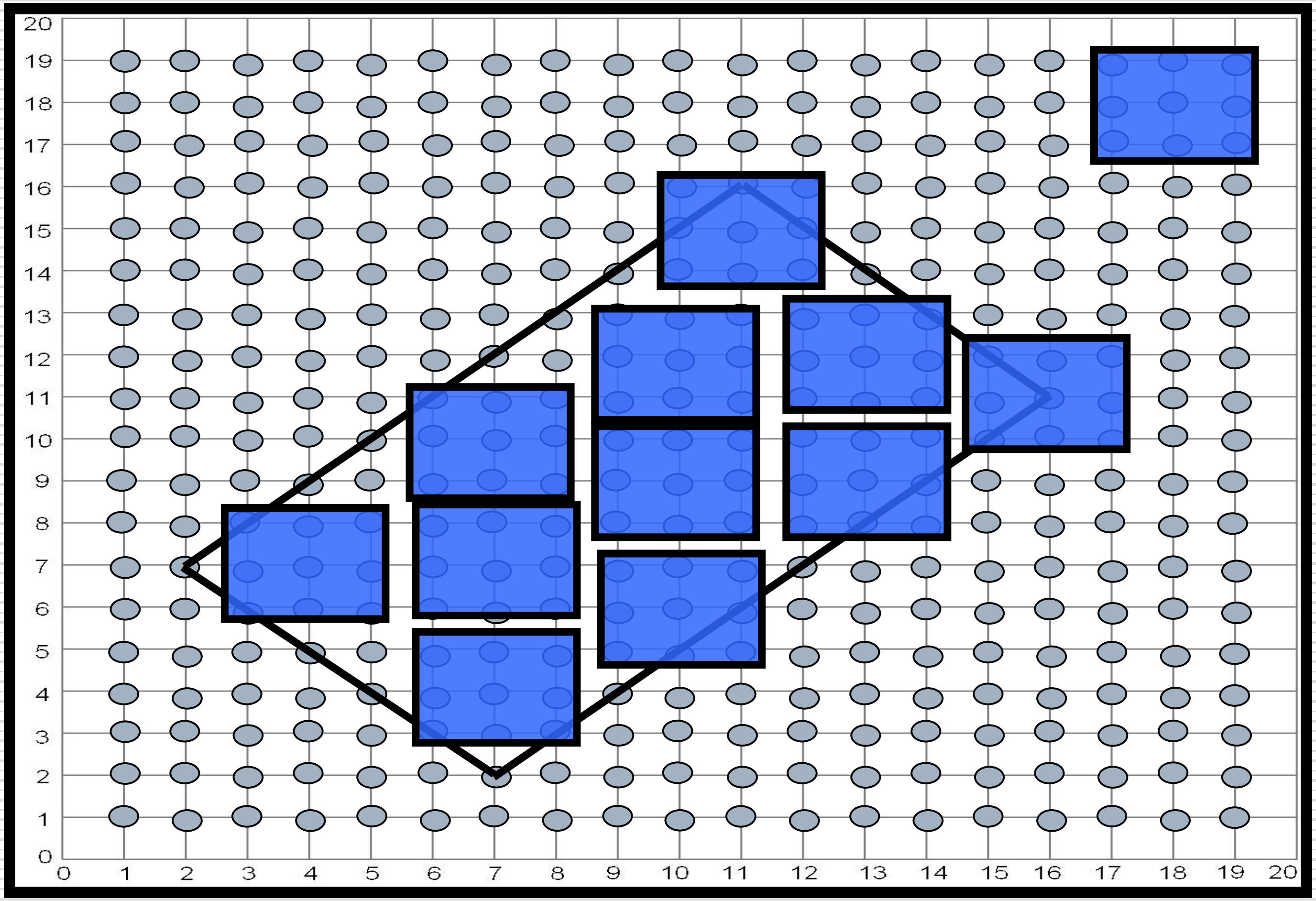
K-Domination

Example Tiling -3 Neighbor Mesh topology



K-Domination

Example Tiling -8 Neighbor Mesh topology



Overview

- Data Aggregation
- K-Domination
  - Definition
  - Embedding functions
  - Tiling in Mesh Topologies
- System Model
  - Energy Model
- Simulation Results
- Conclusions

System Model
Energy Metrics

- Cost of communicating directly to sink is E_s.
- Cost of communicating one hop in the network is e.
- Amount of data compression done at CH is C.

System Model

Energy Metrics

$$K < \frac{\frac{3CEs}{e} - 1}{2}$$

- The above relation gives feasible values of ‘k’ for given energies and compression ratio metrics.
- Value of ‘k’ implies the number of nodes that can be handled by each CH

Overview

- Data Aggregation
- K-Domination
  - Definition
  - Embedding functions
  - Tiling in Mesh Topologies
- System Model
  - Energy Model
- Simulation Results
- Conclusions

Simulation Results

Metrics Used

- Average Latency
  - end-to-end delay averaged over all messages that travel across multi-hop routes
- Average Jitter
  - Variance of average delay over all messages
- Message Loss Rate
  - ratio of number of messages NOT received at cluster head to the number of messages generated in cluster
    - Buffer overflow
    - Link failure

Simulation Results

Simulation Approach

- Modeled multi-hop communication as a multi-stage queuing network
  - Node sent 100 messages of 36 bytes each
  - Nodes queue length is 10
  - Interarrival time between messages was exponentially distributed with varying mean times.
  - Service time in each queue represents
    - Exponentially distributed with mean time.

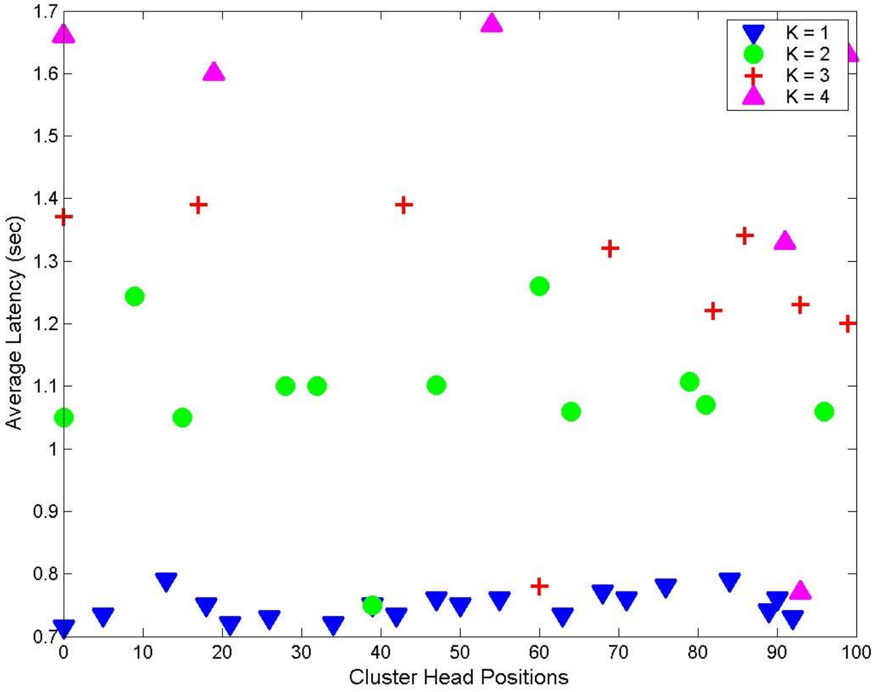
Simulation Results

Metrics Used

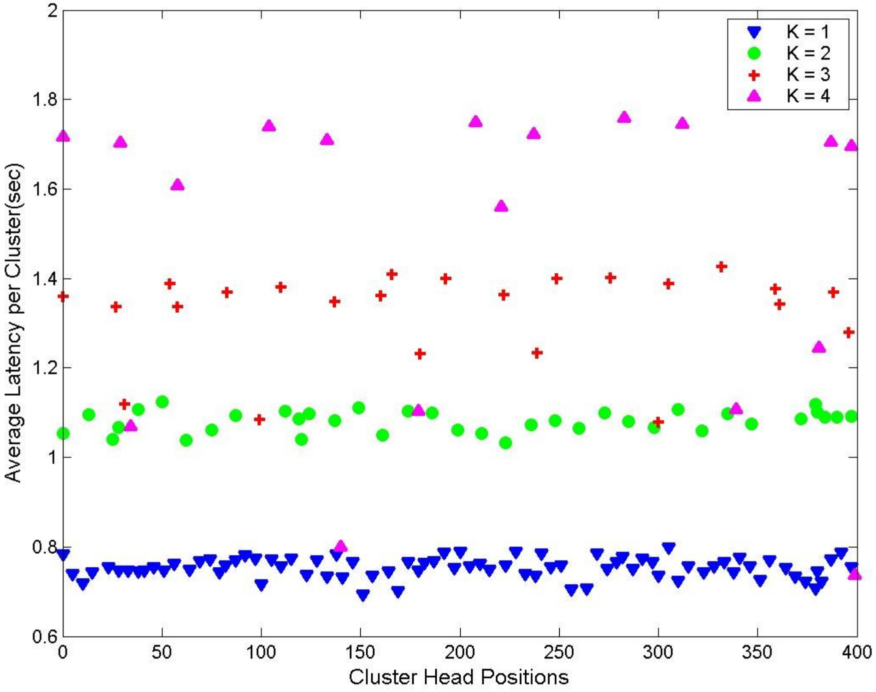
- Average Latency
  - end-to-end delay averaged over all messages that travel across multi-hop routes
- Average Jitter
  - Variance of average delay over all messages
- Message Loss Rate
  - ratio of number of messages NOT received at cluster head to the number of messages generated in cluster
    - Buffer overflow
    - Link failure

Simulation Results

Average Latency

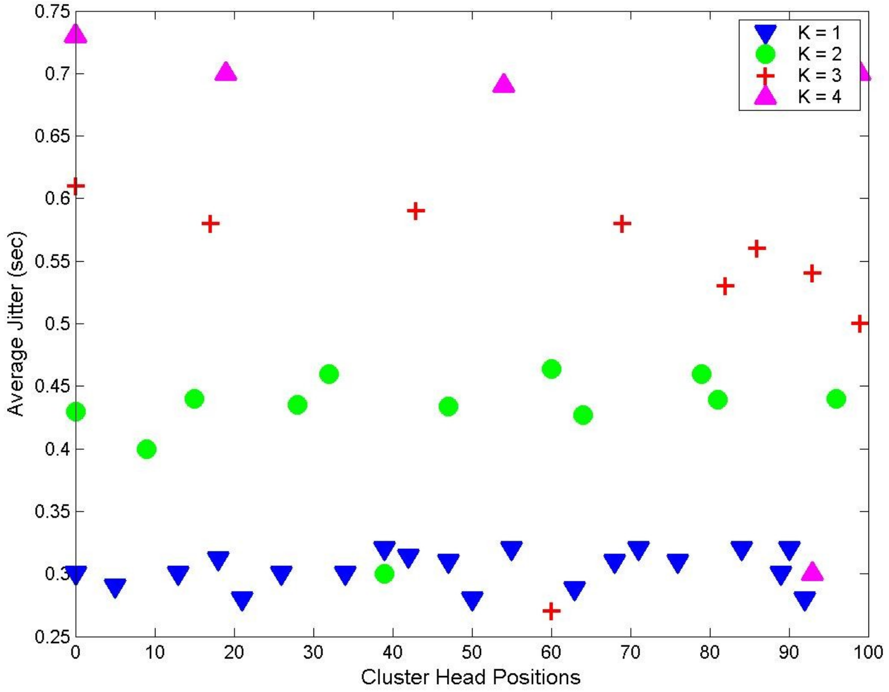


10 x 10 Grid

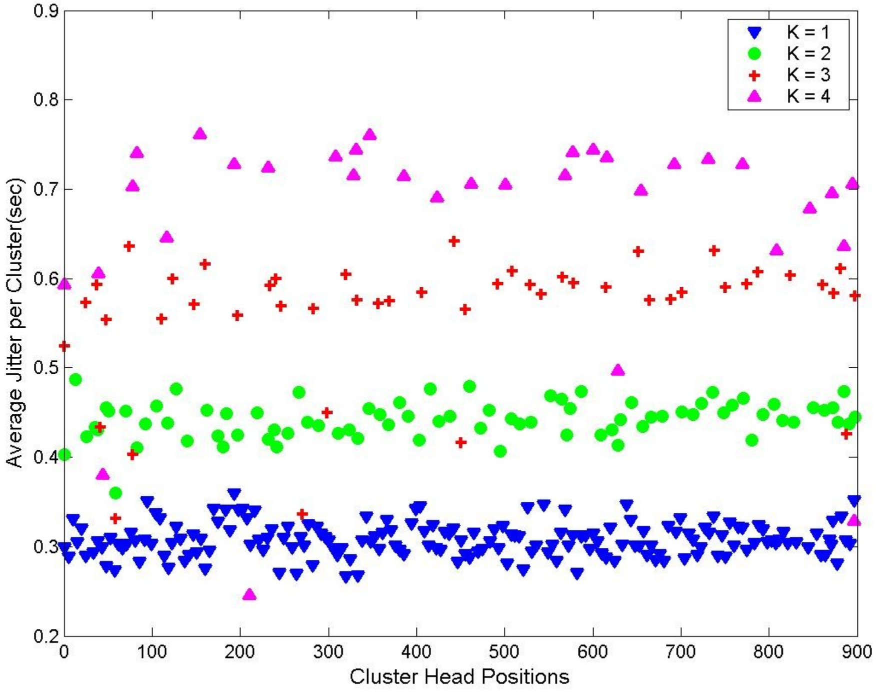


20 x 20 Grid

Simulation Results
Average Jitter

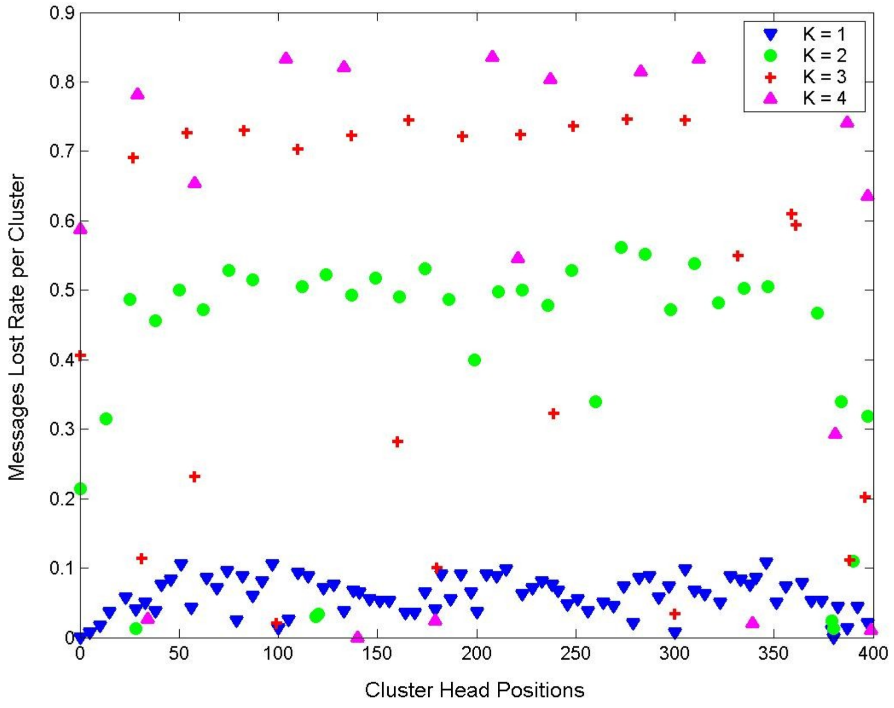


10 x 10 Grid

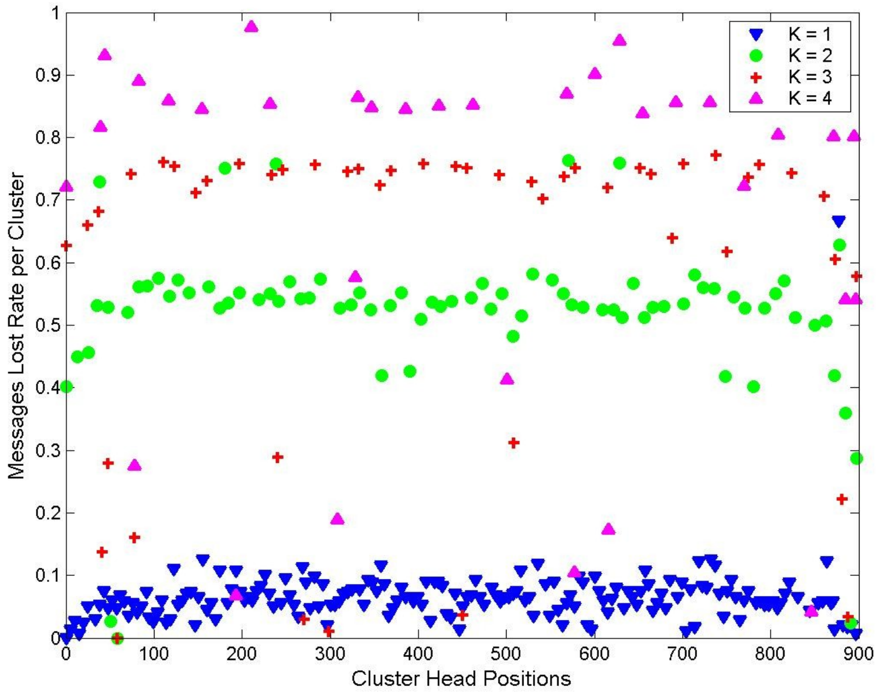


30 x 30 Grid

Simulation Results
Message Loss Rate



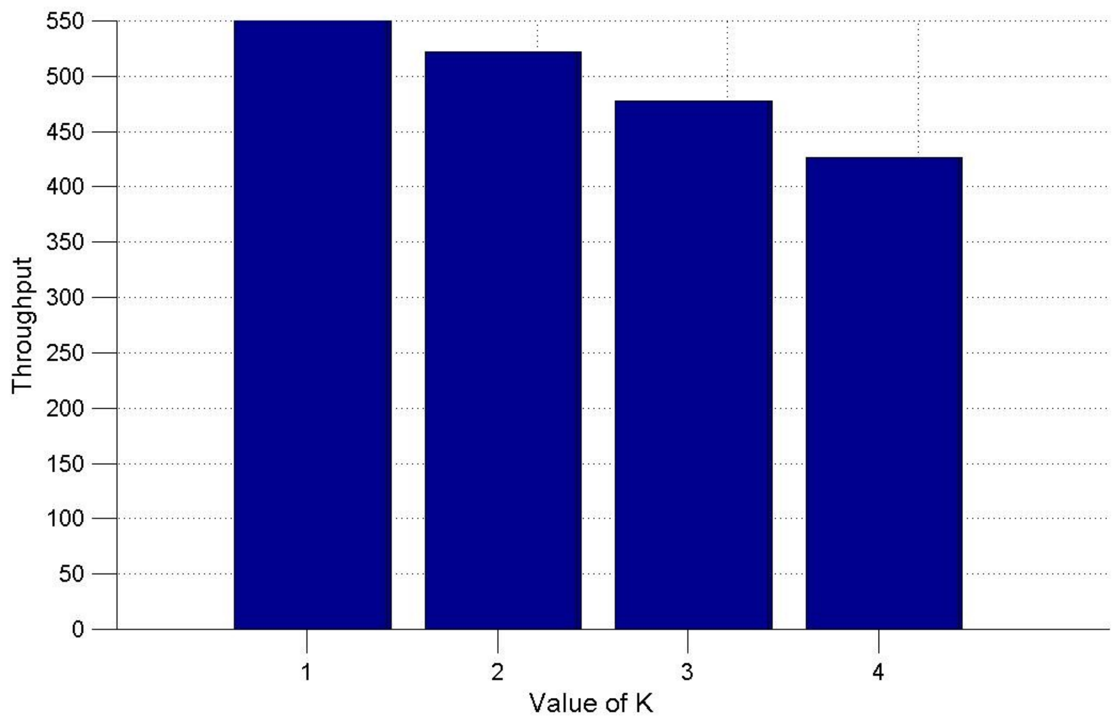
20 x 20 Grid



30 x 30 Grid

Simulation Results

Sink Throughput -30x30 Grid



Overview

- Data Aggregation
- K-Domination
  - Definition
  - Embedding functions
  - Tiling in Mesh Topologies
- System Model
  - Energy Model
- Simulation Results
- Conclusions

Conclusions

- We explored a symmetric way to find the Cluster Heads for a region.
- The cost incurred by nodes to communicate directly is more compared to the cost incurred using CHs.
- The number of CHs effected the performance of the system

Conclusions

- The system performed best when it had less number of CHs
- Future work includes adapting this CH selection scheme for different mesh topologies.
- Cluster head performance analysis based on QoS would provide a clear view.

Research Support

- This research is funded in part by Wright Center for Sensor Systems Engineering (WCSSE) and NSF.

Thank You…

Questions???